

Reviewer's report

Title: Data transfer from animal models to human studies: what is the challenge? Case of persistent organohalogenes.

Version: 1 **Date:** 15 August 2008

Reviewer: William Boyes

Reviewer's report:

This manuscript describes an ambitious assessment of the literature regarding the toxicity of polychlorinated biphenyls (PCBs) and polybrominated diphenyl ethers (PBDEs). The authors pose several important questions regarding the approach to toxicological science, and evaluate trends of publications found in electronic literature searches of Pubmed based on these questions. They then draw conclusions regarding the efficiency (or lack thereof) of toxicological research to inform risk assessments and also the ability of toxicology to provide biomarkers for epidemiological research. The conclusions drawn are rather astounding and, if correct, depressing. Among the conclusions are that there is an average delay of 50 years before regulation of environmental contaminants occurs, and also that there was no learning or improvement in the scientific process from the long history of studying PCBs that transferred to the more recent case of PBDEs. Unfortunately for the manuscript, the evidence presented is not sufficient to support either of these claims. In part this is because the methods for analysis of the scientific literature are not described sufficiently. Largely, however, the authors have not adequately evaluated other plausible explanations for their observations. In addition, some of the suggestions provided by the authors for improving the situation are highly impractical considering the increasingly limited resources being devoted to toxicological assessments.

Major Compulsory Revisions.

1. The claim that an average of 50 years laps before regulation of environmental pollutants is presented as fact without providing the reader with any explanation or documentation. It is not entirely clear where this number came from. I could not find it in any of the references cited. The figure apparently comes from the lag between the introduction of PCBs in the 1920's and the eventual regulation of their use in the 1970s. However, the discovery of the ubiquitous presence of PCBs in fat tissues of humans and many carnivorous animals was key in recognition of their potential danger, and this occurred primarily due to improvements in the sensitivity of analytical chemistry. The ability to measure PCB residues with sensitive analytical methods, in combination with two poisoning episodes, led to the concern that prompted subsequent toxicological studies of PCBs and their regulation. Thus, the "delay" of 50 years had little to do with the pace of toxicological research as the manuscript implies.

2. Also concerning the supposed "50 year" lag before establishing environmental

controls. It is not really valid to consider events beginning in the 1920's as reflecting the current state of affairs. The 1920's, when PCBs were first introduced, preceded the modern environmental movement, the development of toxicological sciences through the mid and late twentieth century, and the establishment of public environmental institutions and environmental legislation in countries around the world. It is not at all clear that an average of 50 years between introduction and control of environmental hazards is currently a valid estimate. Perhaps the authors should take a different approach to address this point, such as looking at recent regulations of environmental hazards passed in Europe, North America, or other developed regions, and assessing the current lag time in controlling those risks. In addition, a representative sample of chemicals should be evaluated, and not a selected set of four compounds, if there is to be any claim that the data are indicative of an average case.

3. The methodology for selecting papers from the literature needs to be described. The literature search identified 6076 PCB articles and 649 PBDE articles, of which only 748 PCB and 59 PBDE articles were analyzed. This reflects only about 10% of the literature identified. How were the papers selected for analysis? Was it random, or were some selection criteria applied? If so, what were they? Were the seven criteria for a "harmonized animal experiment" used in some way to select the papers for analysis?

4. The papers were judged for several criteria to identify "Harmonized animal experiments", but the parameters of these criteria, and how they were used, are not described. How were the criteria applied to the papers selected, or were the papers selected based on the criteria? How was each paper judged to have met, or not met, a criterion? Were they included or excluded from the analysis based on these judgments? Also each of the criteria needs to be more specifically described as to what was necessary for a paper to have met the criterion or not.

5. The first criterion is that the experiments have to be completed in time – i.e. before the substance becomes a public health problem and with enough lead time for public decision making. How was this criteria applied? It would appear that this criterion should eliminate all the research identified on PCBs entirely, since the PCBs were introduced into commerce in the 1920s and electronic literature search did not identify any papers published before 1971.

6. The second criterion refers to doses corresponding to levels of human exposure, but again the specific values for deciding if studies met this criterion are not provided.

7. The authors appear unaware of the "Catch-22" contradiction of the first two criteria. If toxicological studies are conducted in a timely fashion before substance becomes a public health risk, then at that time it cannot be known what dose levels humans will be exposed to.

8. The third criterion for a harmonized animal experiment is that the animal model be similar to humans in toxicokinetics and physiology. Toxicokinetics and physiology are complicated. Again, it is unclear how the authors decided which animals fit this standard and which did not.

9. The fourth and fifth criteria are to study the most sensitive endpoints and most

sensitive stages of development. These things differ across substances, however, making it impossible to start with the most sensitive endpoints and ages at the outset. Not every compound is most toxic to the fetus. Rather the sensitivities typically become known over the course of a series of research projects.

10. The sixth criterion urges the use of endpoints in animal models that can be used as biomarkers of exposure and effect in human epidemiological studies. Yes, this is important. It would be helpful again to understand how this feature was judged by the authors when evaluating the papers.

11. The final criterion concerns estimation of internal doses. Again what were the criteria? Was it necessary to measure tissue concentrations, or were PBPK model estimates sufficient? For measured tissue concentrations, was blood sufficient, or was brain necessary? Was a measure in adults sufficient or did they need to measure in the fetus? Did they need to measure fetal brain concentrations – at what age? What about the dose parameter – peak concentration or area under the curve? What about dose pattern – spikes following i.v. administration vs. more steady concentrations from the diet? What about chemical species / congeners?

12. In the history of PCB and PBDEs provided, it is noted that PBDEs were introduced into commerce in the 1980's and banned in 2004-2008. Thus, the regulatory delay for PBDEs decreased substantially from the 50 year lag for PCBs. This would appear to contradict the authors' conclusion that nothing was learned from the PCB experience.

13. The literature search was confined to the electronic database in PubMed. It should be noted that the electronic databases do not capture well the older scientific literature that predated the internet and existed only in print media. Thus, the first paper identified in 1971 likely reflects the limitations of the electronic database rather than the first toxicological studies of these compounds. Similarly the growing number of publications across the earlier years may reflect the increasing proportion of the published literature that is captured in PubMed as well as the actual trend of research changing over time.

14. Figure 1 illustrates the importance of question #3 above, regarding how the papers were selected. This graph presents the number of papers published / year, but does not include all the 6078 PCB articles or all the 649 PBDE articles found. Apparently, it includes only those papers somehow selected for analysis. Were the articles selected on a random basis, or was the selection stratified by year of publication? If not, then the analysis presented in Figure 1 is not a valid representation of the temporal trend.

15. Figure 2 uses the average daily dose as the dose parameter. Is this measure of dose the most predictive of toxic outcome? It might be that for bio-accumulative and persistent compounds such as the PCBs or PBDEs, the total dose or the area-under-the-curve is more predictive of toxic outcome. In this case, the animal toxicology studies likely would have shorter dose durations than the human exposures, and the cumulative dose measures might have been less different across species. The authors state that the use of a total dose scale did

not change the appearance of the graph. Were the human values also converted to total dose and if so, how? It is also worth considering that peak tissue concentration might be the most important dose parameter determining toxicity. If this were true, then i.v. administration, which provides high dose spikes, will appear to show toxicity at low dose levels if the dose is expressed as the average daily dose.

16. The conclusions section links problems identified in the toxicological literature to delays in regulation without an analysis of the regulatory actions taken and the history of their development. From the analysis presented, therefore, it is not clear that the state of the toxicological sciences was in fact the bottle neck for regulatory development.

17. The lessons learned section of the conclusions provides a number of laudable, but impractical, suggestions such as screening every compound before marketing for developmental neurotoxicity, using the most sensitive endpoints, ages and species. It is already too late to do this now for tens of thousands of existing compounds. In addition, there is not enough money and there are too few laboratory facilities to come close to meeting this suggestion. Expanding the list of species, ages and endpoints assessed will only stretch the fixed resources further, meaning ultimately that fewer (not more) chemicals will be assessed. Surely other approaches should be considered to address this problem.

18. Figures 6, 7, & 8. Some of the plots show a lot of data scatter and a shallow slope of the linear regression. Are the slopes of the linear regression lines significantly different from zero?

Minor Essential Revisions

1. Figure 2. Please move the year-time scale to the bottom of the graph so it does not overlap the plotted data points.
2. The English language grammar and style needs to be improved. The paper should be reviewed by an English language technical writer / editor.
3. There are a number of additional issues with the manuscript that could be raised, but many of them should be fixed by use of an English language editor as suggested above.

Discretionary Revisions

1. The data analyzed for this paper should be made available. This means listing the citations of the papers analyzed and the authors' assessment of each paper on the 7 criteria listed. This is the only way that others could verify the assessment and analyses presented in this paper. Such a data set would obviously be too large to include in a primary publication, but a table with this data could be posted on the internet as supplemental material.

Level of interest: An article of importance in its field

Quality of written English: Not suitable for publication unless extensively edited

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

'I declare that I have no competing interests'